Rational Value of Information Estimation for Measurement Selection

David Tolpin
Computer Science Dept., Ben-Gurion University, 84105 Beer-Sheva, Israel

Solomon Eyal Shimony
Computer Science Dept., Ben-Gurion University, 84105 Beer-Sheva, Israel

ABSTRACT. Computing value of information (VOI) is a crucial task in various aspects of decision-making under uncertainty, such as in meta-reasoning for search; in selecting measurements to make, prior to choosing a course of action; and in managing the exploration vs. exploitation tradeoff. Since such applications typically require numerous VOI computations during a single run, it is essential that VOI be computed efficiently. We examine the issue of anytime estimation of VOI, as frequently it suffices to get a crude estimate of the VOI, thus saving considerable computational resources. As a case study, we examine VOI estimation in the measurement selection problem. Empirical evaluation of the proposed scheme in this domain shows that computational resources can indeed be significantly reduced, at little cost in expected rewards achieved in the overall decision problem.

1 INTRODUCTION

Problems of decision-making under uncertainty frequently contain cases where information can be obtained using some costly actions, called measurement actions. In order to act rationally in the decision-theoretic sense, measurement plans are typically optimized based on some form of value of information (VOI). Computing VOI can also be computationally intensive. Since frequently an exact VOI is not needed in order to proceed (e.g. it is sufficient to determine that the VOI of a certain measurement is much lower than that of another measurement, at a certain point in time), significant computational resources can be saved by controlling the resources used for estimating the VOI. This paper examines this tradeoff via a case study of measurement selection.

In general, computation of value of information (VOI), even under the commonly used simplifying myopic assumption, involves multidimensional integration of a general function [Russell and Wefald, 1991]. For some problems, the integral can be computed efficiently [Russell and Wefald, 1989]; but when the utility function is computationally intensive or when a non-myopic estimate is used, the time required to compute the value of information can be significant [Heckerman et al., 1993] [Bilgic and Getoor, 2007] and must be taken into account while computing the net value of information. This paper presents and analyzes an extension of the known greedy algorithm that decides when to recompute VOI of each of the measurements based on the principles of limited rationality [Russell and Wefald, 1991].

Although it may be possible to use this idea in more general settings, this paper mainly examines on-line most informative measurement selection [Krause and Guestrin, 2007] [Bilgic and Getoor, 2007], an approach which is commonly used to solve problems of optimization under uncertainty [Zheng et al., 2005] [Krause et al., 2008]. Since this approach assumes that the computation time required to select the most informative measurement is negligible compared to the measurement time[Russell and Wefald, 1991], it is important in this setting to ascertain that VOI estimation indeed does not consume excessive computational resources.

2 THE MEASUREMENT SELECTION PROBLEM

As our case study, we examine the following optimization problem. Given:

- A set of N_s items $S = \{s_1, s_2, \dots, s_{N_s}\}$.
- A set of N_f item features $Z = \{z_1, z_2, \dots, z_{N_f}\}$. (Each feature z_i has a domain $\mathcal{D}(z_i)$.)
- A joint distribution over the features of the items in S. That is, a joint distribution over the random variables $\{z_1(s_1), z_2(z_1), \dots, z_1(s_2), z_2(s_2), \dots\}$.
- A set of measurement types $M = \{(c,p)_k \,|\, k \in 1..N_m\}$, with potentially different intrinsic measurement cost c and observation distribution p, conditional on the true feature values, for each measurement type.
- A utility function $u(\mathbf{z})\colon \mathbb{R}^{N_f} \to \mathbb{R}$ on features. In the simplest case, there is just one real-valued feature, acting as the item's utility value, and u is simply the identity function.
- A measurement budget C.

Find a policy of measurement decisions and a final selection that maximize the expected net utility of the selection (the expected reward):

$$\max{:}R = u(\mathbf{z}(s_\alpha)) - \sum_{i=1}^{N_q} c_{k_i} \quad \text{s.t.:} \sum_{i=1}^{N_q} c_{k_i} \le C \tag{1}$$

where $Q = \{(k_i, s_i) \,|\, i \in 1..N_q\}$ is the performed measurement sequence and s_α is the selected item. A next measurement is selected on-line, after the outcomes of all preceding measurements are known.

The above selection problem is intractable, and is therefore commonly solved approximately using a greedy heuristic algorithm. The greedy algorithm selects a measurement $m_{j_{\max}}$ with the greatest net value of information $V_{j_{\max}}$. The *net value of information* is the difference between the intrinsic value of information and the measurement cost.

$$V_j = \Lambda_j - c_{k_j} \tag{2}$$

The *intrinsic value of information* Λ_j is the expected difference in the true utility of the finally selected item s_α after and before the measurement.

$$\Lambda_j = \mathbb{E}(\mathbb{E}[u(\mathbf{z}(z_{\alpha^j}))] - \mathbb{E}[u(\mathbf{z}(s_\alpha))]) \tag{3}$$

Exact computation of Λ_j is intractable, and various estimates are used, including the myopic estimate [Russell and Wefald, 1991] and semi-myopic schemes [Tolpin and Shimony, 2010].

The pseudocode for the algorithm is presented as Algorithm 1. At each step, the algorithm recomputes the value of information estimate of every measurement. The assumptions behind the greedy algorithm are justified when the cost of selecting a next measurement is negligible compared to the measurement cost. However, optimization problems with hundreds and thousands of items are common [Tolpin and Shimony, 2010]; and even if the value of information of a single measurement can be computed efficiently [Russell and Wefald, 1989], the cost of estimating the value of information of all measurements becomes comparable to and outgrows the cost of performing a measurement.

Recomputation of the value of information for every measurement is often unnecessary, especially when using the "blinkered" scheme [Tolpin and Shimony, 2010], a greedy algorithm which attempts to also compute VOI for *sequences* of measurements of the same type. When there are many different

Algorithm 1 Greedy measurement selection

1: $budget \leftarrow C$
2: Initialize beliefs
3: **loop**
4:   **for all** items s_i **do**
5:     Compute $\mathbb{E}(U_i)$
6:   **for all** measurements m_j **do**
7:     **if** $c_j \leq budget$ **then**
8:       Compute V_j
9:     **else**
10:       $V_j \leftarrow 0$
11:   $j_{\max} \leftarrow \arg\max_j V_j$
12:   **if** $V_{j_{\max}} > 0$ **then**
13:     Perform measurement $m_{j_{\max}}$; Update beliefs; $budget \leftarrow budget - c_{j_{\max}}$
14:   **else**
15:     **break**
16: $\alpha \leftarrow \arg\max_j \mathbb{E}(U_i)$
17: **return** s_α

measurements, the value of information of most measurements is unlikely to change abruptly due to just one other measurement results. With an appropriate uncertainty model, it can be shown that the VOI of only a few of the measurements must be recomputed after each measurement, thus decreasing the computation time and ensuring that the greedy algorithm exhibits a more rational behavior w.r.t. computational resources.

3 RATIONAL COMPUTATION OF VALUE OF INFORMATION

For the selective VOI recomputation, the belief $\mathrm{BEL}(\Lambda_j)$ about the intrinsic value of information of measurement m_j is modeled by a normal distribution with variance ς_j^2:

$$\mathrm{BEL}(\Lambda_j) = \mathcal{N}(\Lambda_j, \varsigma_j^2) \tag{4}$$

After a measurement is performed, and the beliefs about the item features are updated (line 13 of Algorithm 1), the belief about Λ_j becomes less certain. Under the assumption that the influence of each measurement on the value of information of other measurements is independent of influence of any other measurement, the uncertainty is expressed by adding Gaussian noise with variance τ^2 to the belief:

$$\varsigma_j^2 \leftarrow \varsigma_j^2 + \tau^2 \tag{5}$$

When Λ_j of measurement m_j is computed, $\mathrm{BEL}(\Lambda_j)$ becomes exact ($\varsigma_j^2 \leftarrow 0$). At the beginning of the algorithm, the beliefs about the intrinsic value of information of measurements are computed from the initial beliefs about item features.

In the algorithm that recomputes the value of information selectively, the initial beliefs about the intrinsic value of information are computed immediately after line 2 in Algorithm 1, and lines 6–11 of Algorithm 1 are replaced by Algorithm 2. While the number of iterations in lines 7–12 of Algorithm 2 is the same as in lines 6–10 of Algorithm 1, W_k is efficiently computable, and the subset of measurements for which the value of information is computed in line 15 of Algorithm 2 is controlled by the computation cost c_V:

$$W_k = \frac{\varsigma_k}{\sqrt{2\pi}} e^{\left(-\frac{(V_\gamma - V_k)^2}{2\varsigma_k^2}\right)} - |V_\gamma - V_k| \, \Phi\left(-\frac{|V_\gamma - V_k|}{\varsigma_k}\right) - c_V \tag{6}$$

Algorithm 2 Rational computation of the value of information

for all measurements m_j **do**
  if $c_j \le budget$ **then**
    $V_j \leftarrow \Lambda_j - c_j; \quad \varsigma_j \leftarrow \sqrt{{\varsigma_j}^2 + \tau^2}$
  else
    $V_j \leftarrow 0; \quad \varsigma_j \leftarrow 0$
loop
  for all measurements m_k **do**
    if $c_k \le budget$ **then**
      Compute W_k
    else
      $W_k \leftarrow 0$
  $k_{\max} \leftarrow \arg\max\limits_k W_k$
  if $W_{k_{\max}} \le 0$ **then**
    break
  Compute $\Lambda_{k_{\max}}; \quad V_{k_{\max}} \leftarrow \Lambda_{k_{\max}} - c_{k_{\max}}; \quad \varsigma_{k_{\max}} \leftarrow 0$
$j_{\max} \leftarrow \arg\max_j V_j$
Compute $\Lambda_{j_{\max}}; \quad V_{j_{\max}} \leftarrow \Lambda_{j_{\max}} - c_{j_{\max}}; \quad \varsigma_{j_{\max}} \leftarrow 0$

where V_γ is the highest value of information V_α if any but the highest value of information is recomputed, and the next to highest value of information V_β if the highest value of information is recomputed; $\Phi(x)$ is the Gaussian cumulative probability of x for $\mu = 0,\ \sigma = 1$.

3.1 Obtaining Uncertainty Parameters

Uncertainty variance τ^2 can be learned as a function of the total cost of performed measurements, either off-line from earlier runs on the same class of problems, or on-line. Learning $\tau^2(c)$ on-line from earlier VOI recomputations proved to be robust and easy to implement: τ^2 is initialized to 0 and gradually updated with each recomputation of the value of information.

4 EMPIRICAL EVALUATION

Experiments in this section compare performance of the algorithm that recomputes the value of information selectively with the original algorithm in which the value of information of every measurement is recomputed at every step. Two of the problems evaluated in [Tolpin and Shimony, 2010] are considered: *noisy Ackley function maximization* and *SVM parameter search*. For each of the optimization problems, plots of the number of VOI recomputations, the reward, the intrinsic utility, and the total cost of measurements are presented. The results are averaged for multiple (100) runs of each experiment, such that the standard deviation of the reward is $\approx 5\%$ of the mean reward. In the plots, the solid line corresponds to the rationally recomputing algorithm, the dashed line corresponds to the original algorithm, and the dotted line corresponds to the algorithm that selects measurements randomly and performs the same number of measurements as the rationally recomputing algorithm for the given computation cost c_V. Since, as can be derived from (6), the computation time T_r of the rationally recomputing algorithm decreases with the logarithm of the computation cost c_V, $T_r = \Theta(A - B\log c_V)$, the computation cost axis is scaled logarithmically.

4.1 The Ackley Function

The Ackley function [Ackley, 1987] is a popular optimization benchmark. The two-argument form of the Ackley function is used in the experiment; the function is defined by the expression (7):

$$A(x,y) = 20 \cdot \exp\left(-0.2\sqrt{\frac{x^2+y^2}{2}}\right) + \exp\left(\frac{cos(2\pi x)+cos(2\pi y)}{2}\right) \tag{7}$$

In the optimization problem, the utility function is $u(z) = \tanh(2z)$, the measurements are normally distributed around the true values with variance $\sigma_m^2 = 0.5$, and the measurement cost is 0.01. There are uniform dependencies with $\sigma_w^2 = 0.5$ in both directions of the coordinate grid with a step of 0.2 along each axis. The results for the blinkered scheme[Tolpin and Shimony, 2010] are presented in

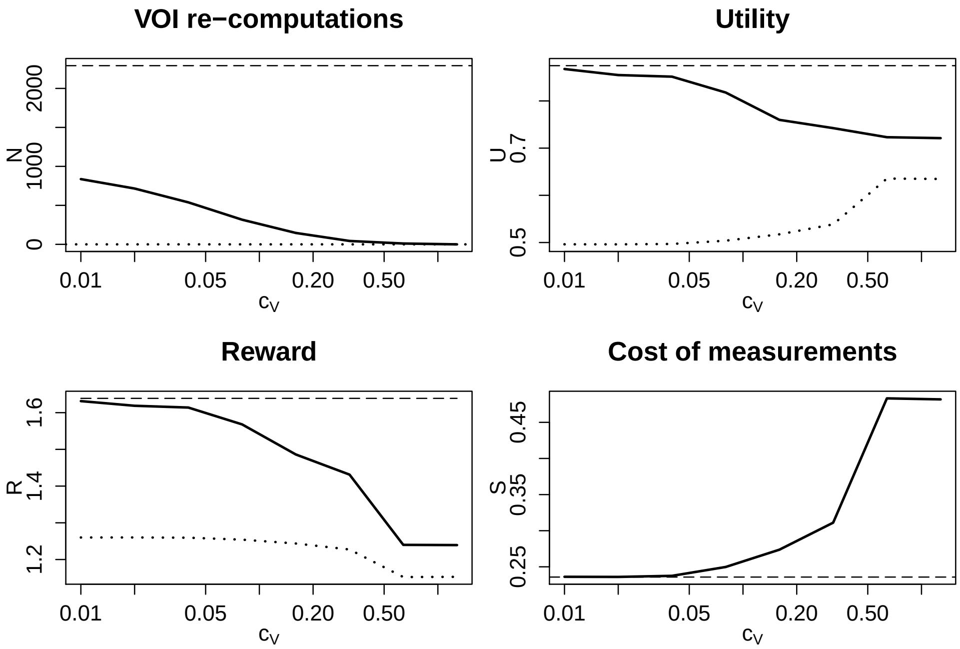

Figure 1: The Ackley function, blinkered scheme.

Figure 1.

4.2 SVM Parameter Search

An SVM (Support Vector Machine) classifier based on the radial basis function has two parameters: C and γ. A combination of C and γ with high expected classification accuracy should be chosen, and an efficient algorithm for determining the optimal values is not known. A trial for a combination of parameters determines estimated accuracy of the classifier through cross-validation. The SVMGUIDE2 [wei Hsu et al., 2003] dataset is used for the case study. The utility function is $u(z) = \tanh(4(z-0.5))$, the $\log C$ and $\log\gamma$ axes are scaled for uniformity to ranges [1..21] and there are uniform dependencies along both axes with $\sigma_w^2 = 0.4$. The measurements are normally distributed with variance $\sigma_m^2 = 0.25$ around the true values, and the measurement cost is $c_m = 0.01$. The results for the myopic scheme are presented in Figure 2.

4.3 Discussion of Results

In all experiments, a significant decrease in the computation time is achieved with only slight degradation of the reward; performance of the rationally recomputing algorithm decreases slowly with the computation cost and exceeds performance of the algorithm that makes random measurements even when VOI for only a small fraction of measurements is recomputed at each step.

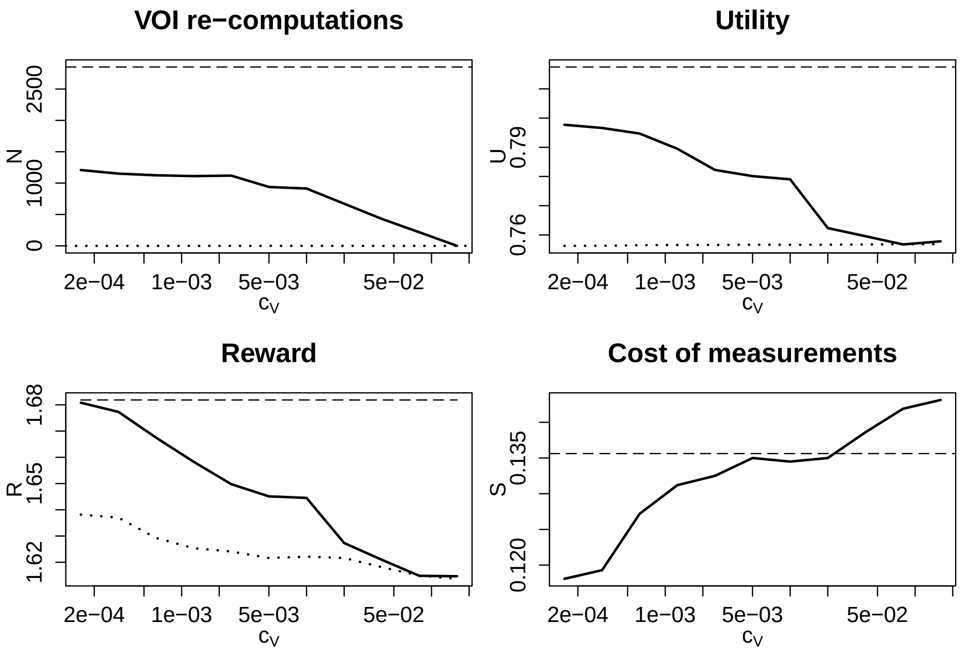

Figure 2: SVM parameter search, myopic scheme.

Exact dependency of performance of the rationally recomputing of algorithm on the intensity of VOI recomputations varies among problems and depends both on the problem properties and on the VOI estimate used in the algorithm.

5 CONCLUSION

The paper proposes an improvement to a widely used class of VOI-based optimization algorithms. The improvement allows to decrease the computation time while only slightly affecting the performance. The proposed algorithm rationally reuses computations of VOI and recomputes VOI only for measurements for which a change in VOI is likely to affect the choice of the next measurement.

6 ACKNOWLEDGEMENTS

The research is partially supported by the IMG4 Consortium under the MAGNET program of the Israeli Ministry of Trade and Industry, by Israel Science Foundation grant 305/09, by the Lynn and William Frankel Center for Computer Sciences, and by the Paul Ivanier Center for Robotics Research and Production Management.

References

[Ackley, 1987] Ackley, D. H. (1987). *A connectionist machine for genetic hillclimbing.* Kluwer Academic Publishers, Norwell, MA, USA.

[Bilgic and Getoor, 2007] Bilgic, M. and Getoor, L. (2007). Voila: Efficient feature-value acquisition for classification. In *AAAI*, pages 1225–1230. AAAI Press.

[Heckerman et al., 1993] Heckerman, D., Horvitz, E., and Middleton, B. (1993). An approximate nonmyopic computation for value of information. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(3):292–298.

[Krause and Guestrin, 2007] Krause, A. and Guestrin, C. (2007). Near-optimal observation selection using submodular functions. In *AAAI*, pages 1650–1654.

[Krause et al., 2008] Krause, A., Leskovec, J., Guestrin, C., VanBriesen, J., and Faloutsos, C. (2008). Efficient sensor placement optimization for securing large water distribution networks. *Journal of Water Resources Planning and Management*, 134(6):516–526. (Draft; full version available here).

[Russell and Wefald, 1989] Russell, S. J. and Wefald, E. (1989). On optimal game-tree search using rational meta-reasoning. In *IJCAI*, pages 334–340.

[Russell and Wefald, 1991] Russell, S. J. and Wefald, E. (1991). *Do the right thing: studies in limited rationality*. MIT Press, Cambridge, MA, USA.

[Tolpin and Shimony, 2010] Tolpin, D. and Shimony, S. E. (2010). Semi-myopic measurement selection for optimization under uncertainty. Technical Report 10-01, Lynne and William Frankel Center for Computer Science at Ben Gurion University of the Negev, Israel.

[wei Hsu et al., 2003] wei Hsu, C., chung Chang, C., and jen Lin, C. (2003). A practical guide to support vector classification. Technical report.

[Zheng et al., 2005] Zheng, A. X., Rish, I., and Beygelzimer, A. (2005). Efficient test selection in active diagnosis via entropy approximation. In *UAI*, pages 675–682.